GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:47:20 ; Search time 210.42 Seconds
(without alignments)
7.638 Million cell updates/sec

Title: US-09-331-631A-8_COPY_33_79
Perfect score: 275
Sequence: 1 GDDDPPKRYEDCRRRCEWDT.....QCEESCKSQYGEKDQQQRHR 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	275	100.0	590	19	W62832	Gossypium hirsutum
2	124	45.1	525	19	W62831	Theobroma cacao an
3	124	45.1	566	13	R20181	Sequence encoded b
4	94	34.2	666	19	W62829	Macadamia integrif
5	92	33.5	666	19	W62828	Macadamia integrif
6	88	32.0	625	19	W62830	Macadamia integrif
7	62	22.5	360	20	Y33492	Human Machado-Jose
8	61	22.2	28	19	W62841	Stenocarpus sinuat
9	61	22.2	265	12	R12844	HTLV-1 protein exp
10	60	21.8	259	20	W90009	Expressed antigen
11	60	21.8	303	15	R60054	Dirofilaria immiti
12	60	21.8	810	20	Y35710	Chlamydia pneumoni
13	60	21.8	810	20	Y34602	Chlamydia pneumoni
14	60	21.8	1743	19	W98879	H. pylori GHPO 175
15	59.5	21.6	593	19	W62835	Zea mays antimicro
16	59	21.5	491	18	W35292	Human disintegrin
17	59	21.5	691	21	Y79033	Human Kuz amino ac
18	59	21.5	748	19	W70457	Mutant human disin
19	59	21.5	748	19	W56132	Homo sapiens trans
20	59	21.5	748	20	Y16776	Human disintegrin
21	59	21.5	799	18	W35293	Human disintegrin
22	59	21.5	799	19	W70456	Human disintegrin
23	59	21.5	2150	21	Y53898	Amino acid sequenc
24	58	21.1	71	21	Y74316	Neisseria meningit
25	58	21.1	84	21	Y74315	Neisseria meningit
26	57.5	20.9	162	20	Y30436	Mature nematode ex
27	57.5	20.9	181	17	R91711	AcaNAP45. Ancylos
28	57.5	20.9	181	20	Y30409	Nematode extracted
29	57.5	20.9	443	18	W32110	Human extracellula
30	57.5	20.9	443	20	Y16587	Extracellular prot
31	57.5	20.9	443	21	Y84706	Amino acid sequenc
32	57.5	20.9	443	21	Y84707	A human p53 mutant
33	57.5	20.9	443	21	Y55850	Human S1-5 ECMP-li
34	57.5	20.9	671	21	Y99426	Human PRO1604 (UNQ
35	57	20.7	342	20	Y16785	Human secreted pro
36	57	20.7	1898	20	Y30795	A human trichohyal
37	56.5	20.5	174	20	W89904	Antigen from clust
38	56.5	20.5	432	20	W93954	Human regulatory m
39	56	20.4	267	20	W89907	Antigen from clust
40	55.5	20.2	434	17	R96419	Peptide fragment o
41	55.5	20.2	434	17	R96420	Peptide fragment o
42	55.5	20.2	2237	14	R33550	Sequence of the al
43	55.5	20.2	2237	16	R71006	Human neuronal cal
44	55.5	20.2	2237	19	W63142	Human calcium chan
45	55.5	20.2	2337	19	W37878	Human calcium chan

ALIGNMENTS

RESULT 1
W62832
ID W62832 standard; Protein; 590 AA.
XX
AC W62832;
XX
DT 27-OCT-1998 (first entry)
XX
DE Gossypium hirsutum antimicrobial protein.
XX
KW antimicrobial protein; infestation; control.
XX
OS Gossypium hirsutum.
XX
PN WO9827805-A1.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-AU00874.
XX
PR 20-DEC-1996; 96AU-0004275.
XX
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
DR WPI; 1998-377279/32.
XX
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT useful for controlling microbial infestations of plants or mammals
XX
PS Claim 1; Page 49-51; 96pp; English.
XX
CC The sequence is that of an antimicrobial protein which can
CC be used to control microbial infestations in plants and mammalian

```
CC   animals.
XX
SQ   Sequence   590 AA;


  Query Match             100.0%;  Score 275;  DB 19;  Length 590;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-23;
  Matches   47;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GDDDPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQRHR 47
           |||||||||||||||||||||||||||||||||||||||||||||||
Db      33 gdddppkryedcrrrcewdtrgqkeqqqceescksqygekdqqqrhr 79


RESULT   2
W62831
ID   W62831 standard; Protein; 525 AA.
XX
AC   W62831;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Theobroma cacao antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Theobroma cacao.
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 47-49; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   525 AA;


  Query Match             45.1%;  Score 124;  DB 19;  Length 525;
  Best Local Similarity   46.3%;  Pred. No. 2.1e-06;
  Matches   19;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 DPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQ 44
           || ::|| |:|||| :   ::||:|||: |: :| |: :||
Db      37 dprqqyeqcqrrceseateereqeqceqrcereykeqqrqq 77


RESULT   3
R20181
ID   R20181 standard; Protein; 566 AA.
XX
AC   R20181;
XX
DT   16-APR-1992  (first entry)
XX
DE   Sequence encoded by 67 kD T. cacao protein cDNA.
XX
KW   Cocoa; flavour; vicilin; seed storage protein.
XX
OS   Theobroma cacao.
XX
PN   WO9119801-A.
XX
PD   26-DEC-1991.
XX
PF   07-JUN-1991;   91WO-GB00914.
XX
PR   11-JUN-1990;   90GB-0013016.
XX
PA   (MRSC ) MARS UK LTD.
XX
PI   Spencer ME,  Hodge R,  Deakin EA,  Ashton S;
XX
DR   WPI; 1992-024418/03.
DR   N-PSDB; Q20377.
XX
PT   Recombinant cocoa proteins - are responsible for flavour in cocoa
PT   beans and produced in large quantities using yeast and bacterial
PT   expression vectors
XX
PS   Claim 4; Fig 2; 59pp; English.
XX
CC   The inventors claim a 67 kD and 31 kD T. cacao protein, and
CC   fragments, and encoding DNAs.  The 47 kD and 31 kD proteins are
CC   derived from the 67 kD precursor.  T. cacao protein cDNA was
CC   detected in a cDNA library prepared from immature cocoa beans RNA
CC   using a probe based on the AA sequence of a CNBr peptide common to
CC   the 47 kD and 31 kD polypeptides.  Homology searches revealed close
CC   homologies between the 67 kD polypeptide and the vicilins, which are
CC   seed storage proteins.
XX
SQ   Sequence   566 AA;


  Query Match             45.1%;  Score 124;  DB 13;  Length 566;
  Best Local Similarity   46.3%;  Pred. No. 2.2e-06;
  Matches   19;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 DPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQ 44
           || ::|| |:|||| :   ::||:|||: |: :| |: :||
Db      37 dprqqyeqcqrrceseateereqeqceqrcereykeqqrqq 77


RESULT   4
W62829
ID   W62829 standard; Protein; 666 AA.
XX
AC   W62829;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Macadamia integrifolia antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Macadamia integrifolia.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..28
FT                   /note= "signal peptide"
FT   Protein         29..666
FT                   /note= "mature protein"
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
```

```
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
DR   N-PSDB; V42311.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 39-41; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   666 AA;


  Query Match             34.2%;  Score 94;  DB 19;  Length 666;
  Best Local Similarity   35.6%;  Pred. No. 0.0058;
  Matches   16;  Conservative    8;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       3 DDPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQRHR 47
           |||    : |:|||     | ::|| |:  ||    |::: | |
Db      76 ddpqtdcqqcqrrcrqqesgprqqqycqrrckeiceeeeeynrqr 120


RESULT   5
W62828
ID   W62828 standard; Protein; 666 AA.
XX
AC   W62828;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Macadamia integrifolia antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Macadamia integrifolia.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..28
FT                   /note= "signal peptide"
FT   Protein         29..666
FT                   /note= "mature protein"
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
DR   N-PSDB; V42310.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 34-36; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   666 AA;


  Query Match             33.5%;  Score 92;  DB 19;  Length 666;
  Best Local Similarity   33.3%;  Pred. No. 0.0097;
  Matches   15;  Conservative    9;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       3 DDPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQRHR 47
           :||    : |:|||     | ::|| |:  ||    |::: | |
Db      76 edpqtecqqcqrrcrqqesgprqqqycqrrckeiceeeeeynrqr 120


RESULT   6
W62830
ID   W62830 standard; Protein; 625 AA.
XX
AC   W62830;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Macadamia integrifolia antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Macadamia integrifolia.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..28
FT                   /note= "signal peptide"
FT   Protein         29..666
FT                   /note= "mature protein"
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
DR   N-PSDB; V42316.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 43-45; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   625 AA;


  Query Match             32.0%;  Score 88;  DB 19;  Length 625;
  Best Local Similarity   29.5%;  Pred. No. 0.025;
  Matches   13;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 DPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQRHR 47
           || ::|| |::||:      :  | |:: |: :| :: ::|: |
Db      80 dpqqqyeqcqkrcqrreteprhmqicqqrcerryekekrkqqkr 123


RESULT   7
Y33492
ID   Y33492 standard; Protein; 360 AA.
XX
AC   Y33492;
```

```
XX
DT   19-JAN-2000  (first entry)
XX
DE   Human Machado-Joseph disease associated protein.
XX
KW   Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;
KW   huntingtin polypeptide; Machado-Joseph disease; SCA1; SCA2; SCA6;
KW   atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;
KW   Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;
KW   dentatorubropallidoluysian atrophy; cell proliferation; cell survival;
KW   neoplastic; malignant; autoimmune; fibrotic.
XX
OS   Homo sapiens.
XX
PN   WO9945944-A1.
XX
PD   16-SEP-1999.
XX
PF   11-MAR-1999;   99WO-US05250.
XX
PR   12-MAR-1998;   98US-0041886.
XX
PA   (BURN-) BURNHAM INST.
XX
PI   Bredesen DE,  Rabizadeh S;
XX
DR   WPI; 1999-561617/47.
DR   N-PSDB; Z23425.
XX
PT   New proapoptotic dependence peptides, used to develop products for
PT   treating, e.g. Alzheimer's disease -
XX
PS   Disclosure; Page 95-96; 199pp; English.
XX
CC   This invention describes novel pure proapoptotic dependence peptides
CC   which comprise a sequence of an active dependence domain selected from
CC   dependence polypeptides consisting of p75NTR, androgen receptor, DCC,
CC   huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2,
CC   SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable
CC   of inducing cell death and can be used to develop products to mediate or
CC   inhibit apoptosis. The methods can be used for reducing the severity of
CC   a proapoptotic dependence domain mediated pathological conditions e.g.
CC   Huntington's disease, Alzheimer's disease, Kennedy's disease,
CC   Spinocerebellar ataxias, dentatorubropallidoluysian atrophy,
CC   Machado-Joseph disease, stroke or head trauma. They can also be used for
CC   reducing the severity of a pathological condition mediated by upregulated
CC   cell proliferation or cell survival e.g. neoplastic, malignant,
CC   autoimmune or fibrotic conditions. This sequence represents a human
CC   Machado-Joseph disease associated protein described in the method of the
CC   invention.
XX
SQ   Sequence   360 AA;


  Query Match             22.5%;  Score 62;  DB 20;  Length 360;
  Best Local Similarity   35.0%;  Pred. No. 12;
  Matches   14;  Conservative   12;  Mismatches   12;  Indels    2;  Gaps    1;

Qy      10 EDCRRRCE--WDTRGQKEQQQCEESCKSQYGEKDQQQRHR 47
           |: |:| |  :: : ||:||| ::  : |  :: |||: |
Db     279 eelrkrreayfekqqqkqqqqqqqqqqqqqqqqqqqqqqr 318


RESULT   8
W62841
ID   W62841 standard; Protein; 28 AA.
XX
AC   W62841;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Stenocarpus sinuatus antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Stenocarpus sinuatus.
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 66; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   28 AA;


  Query Match             22.2%;  Score 61;  DB 19;  Length 28;
  Best Local Similarity   33.3%;  Pred. No. 1.2;
  Matches    9;  Conservative    9;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       4 DPPKRYEDCRRRCEWDTRGQKEQQQCE 30
           || :: : |: ||:   :  ::|||||:
Db       2 dpirqqqlcqmrcqqqekdprqqqqck 28


RESULT   9
R12844
ID   R12844 standard; Protein; 265 AA.
XX
AC   R12844;
XX
DT   18-SEP-1991  (first entry)
XX
DE   HTLV-1 protein expressed by antisense nucleotides 1589-2383.
XX
KW   human T-cell lymphotrophic virus-1; ORF; vaccine; antisense RNA.
XX
OS   Human T cell lymphotrophic virus.
XX
PN   US4999421-A.
XX
PD   12-MAR-1991.
XX
PF   27-JUN-1988;   88US-0211749.
XX
PR   27-JUN-1988;   88US-0211749.
XX
PA   (TRIT-) TRITON BIOSCIENCES.
XX
PI   Brunck TK,  Larocca DJ,  Monahan JJ;
XX
DR   WPI; 1991-206841/28.
DR   N-PSDB; Q12502.
XX
PT   Proteins encoded by ribonucleic acid anti-sense strand - to human
PT   T-cell leukaemia virus 1 viral RNA useful as diagnostic agents,
PT   vaccines and therapeutic agents
XX
PS   Claim 1; Column 16; 15pp; English.
```

```
XX
CC   This polypeptide is encoded by nucleotides 1589-2383 in the reverse
CC   complement of the HTLV-I genome (complementary to nucleotides 7479-
CC   6685, 3'-5' of the HTLV-I sense mRNA). Peptide fragments
CC   corresponding to at least 10 amino acids from this sequence are
CC   disclosed. See also Q12499-Q12501 and Q12503.
XX
SQ   Sequence   265 AA;


  Query Match             22.2%;  Score 61;  DB 12;  Length 265;
  Best Local Similarity   26.5%;  Pred. No. 11;
  Matches   13;  Conservative   17;  Mismatches   17;  Indels    2;  Gaps    1;

Qy       1 GDDDPPK--RYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQRHR 47
           |:  ||:   : | :|| |   : :||::: ||   ::| :: :::: |
Db     123 gekapprgethrdrqrraeekrkrkkerekeeekqtaeylkrkeeekar 171


RESULT  10
W90009
ID   W90009 standard; Protein; 259 AA.
XX
AC   W90009;
XX
DT   18-FEB-1999  (first entry)
XX
DE   Expressed antigen for cluster 33.
XX
KW   Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW   peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
XX
OS   Helicobacter pylori.
XX
PN   WO9849314-A2.
XX
PD   05-NOV-1998.
XX
PF   27-APR-1998;   98WO-US08487.
XX
PR   14-OCT-1997;   97US-0061958.
PR   25-APR-1997;   97US-0045107.
XX
PA   (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI   Chow TP,  Fry KE,  Lim MY,  McAtee CP;
XX
DR   WPI; 1999-009433/01.
XX
PT   New Helicobacter pylori antigens and related nucleic acid sequences
PT   - useful in serological diagnosis and protective vaccines, providing
PT   long-lasting immune response
XX
PS   Claim 16; Page 336-337; 402pp; English.
XX
CC   The present sequence represents a Helicobacter pylori antigenic protein
CC   that is characterised by immunoreactivity with H. pylori-positive
CC   antisera. The proteins are highly immunogenic and induce a long-lasting
CC   immune response that persists even after antimicriobial treatment. In
CC   antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC   highly sensitive and specific. The specification also describes 69
CC   previously unrecognised immunogenic cluster families. H. pylori antigens
CC   are used to detect H. pylori-specific antibodies, for diagnosing
CC   infection or to confirm eradication of infection, and in vaccines to
CC   protect against H. pylori infection and related diseases (gastritis,
CC   peptic ulcer, gastric adenocarcinoma/lymphoma).
XX
SQ   Sequence   259 AA;


  Query Match             21.8%;  Score 60;  DB 20;  Length 259;
  Best Local Similarity   34.2%;  Pred. No. 14;
  Matches   13;  Conservative    9;  Mismatches   12;  Indels    4;  Gaps    1;

Qy       7 KRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQ 44
           | |:||  :     | :||:|:||:    :  :| :||
Db     173 kaykdcvsk----arnekekqecekllltpearkkleqq 206


RESULT  11
R60054
ID   R60054 standard; Protein; 303 AA.
XX
AC   R60054;
XX
DT   21-MAR-1995  (first entry)
XX
DE   Dirofilaria immitis parasitic helminth protein.
XX
KW   Clones; p4; p22U; infection; heartworm; vaccine; larval.
XX
OS   Dirofilaria immitis.
XX
PN   WO9415593-A.
XX
PD   21-JUL-1994.
XX
PF   12-JAN-1994;   94WO-US00679.
XX
PR   12-JAN-1993;   93US-0003257.
PR   12-JAN-1993;   93US-0003389.
PR   19-AUG-1993;   93US-0109391.
XX
PA   (PARA-) PARAVAX INC.
PA   (COLS ) UNIV COLORADO STATE RES FOUND.
XX
PI   Frank GR,  Grieve RB,  Mika-grieve M,  Tripp CA;
XX
DR   WPI; 1994-248861/30.
DR   N-PSDB; Q70048.
XX
PT   New parasitic helminth proteins, pref. from Dirofilaria immitis -
PT   used to protect animals from helminth infection, pref. heartworm
XX
PS   Disclosure; Page 121; 154pp; English.
XX
CC   The sequence is that of the protein product of larval Dirofilaria
CC   immitis clone p4.  The parasitic helminth protein  can
CC   be used to protect animals against helminth infection, esp.
CC   heartworm.
CC   See also R60055.
XX
SQ   Sequence   303 AA;


  Query Match             21.8%;  Score 60;  DB 15;  Length 303;
  Best Local Similarity   37.5%;  Pred. No. 16;
  Matches   15;  Conservative   12;  Mismatches   11;  Indels    2;  Gaps    2;

Qy       8 RYEDCRRRCEWDTRGQKEQQQCE-ESCKSQYGEKDQQQRH 46
           | |: | | | : | |||::: | |  : :| ||::|:::
Db     262 rqeeererqeqerr-qkerermeqerirqeydekerqrqy 300


RESULT  12
Y35710
ID   Y35710 standard; Protein; 810 AA.
XX
AC   Y35710;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Chlamydia pneumoniae transmembrane protein sequence.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
```

```
KW   vaccine; neutralising epitope.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 1415-1416; Disclosure; 1912pp; English.
XX
CC   Y34584-Y35879 represent the proteins encoded by all the open reading
CC   frames in the complete genome (see X91990) of Chlamydia pneumoniae.
CC   C. pneumoniae causes respiratory disease such as pneumonia and
CC   bronchitis and is thought to be a contributing factor in heart
CC   disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC   nodosum or pharyngitis. The polypeptides encoded by the open reading
CC   frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
CC   immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC   nucleotides sequences can also be used as immunogenic compositions,
CC   especially where the vector directs the expression of a neutralising
CC   epitope of C. pneumoniae.
XX
SQ   Sequence   810 AA;


  Query Match             21.8%;  Score 60;  DB 20;  Length 810;
  Best Local Similarity   26.2%;  Pred. No. 44;
  Matches   11;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 DDDPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQ 43
           |:|  : | :|::| : |:  : | : | |  : :  | : |
Db     620 dedlrraytecqkrfqgdsglesevracreqlreriqefetq 661


RESULT  13
Y34602
ID   Y34602 standard; Protein; 810 AA.
XX
AC   Y34602;
XX
DT   13-SEP-1999  (first entry)
XX
DE   Chlamydia pneumoniae transmembrane protein sequence.
XX
KW   Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW   sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW   vaccine; neutralising epitope.
XX
OS   Chlamydia pneumoniae.
XX
PN   WO9927105-A2.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-IB01890.
XX
PR   04-NOV-1998;   98US-0107078.
PR   21-NOV-1997;   97FR-0014673.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-357842/30.
XX
PT   Genome sequence of Chlamydia pneumoniae
XX
PS   Page 625-627; Disclosure; 1912pp; English.
XX
CC   Y34584-Y35879 represent the proteins encoded by all the open reading
CC   frames in the complete genome (see X91990) of Chlamydia pneumoniae.
CC   C. pneumoniae causes respiratory disease such as pneumonia and
CC   bronchitis and is thought to be a contributing factor in heart
CC   disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC   nodosum or pharyngitis. The polypeptides encoded by the open reading
CC   frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
CC   immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC   nucleotides sequences can also be used as immunogenic compositions,
CC   especially where the vector directs the expression of a neutralising
CC   epitope of C. pneumoniae.
XX
SQ   Sequence   810 AA;


  Query Match             21.8%;  Score 60;  DB 20;  Length 810;
  Best Local Similarity   26.2%;  Pred. No. 44;
  Matches   11;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 DDDPPKRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQ 43
           |:|  : | :|::| : |:  : | : | |  : :  | : |
Db     620 dedlrraytecqkrfqgdsglesevracreqlreriqefetq 661


RESULT  14
W98879
ID   W98879 standard; Protein; 1743 AA.
XX
AC   W98879;
XX
DT   31-MAR-1999  (first entry)
XX
DE   H. pylori GHPO 1755 protein.
XX
KW   GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW   peptic ulcer disease.
XX
OS   Helicobacter pylori.
XX
PN   WO9843478-A1.
XX
PD   08-OCT-1998.
XX
PF   01-APR-1998;   98WO-US06371.
XX
PR   29-JUL-1997;   97US-0902615.
PR   01-APR-1997;   97US-0833457.
PR   24-JUN-1997;   97US-0881227.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
PI   Al-Garawi A,  Kleanthous H,  Miller C,  Oomen RP,  Tomb J;
XX
DR   WPI; 1998-542293/46.
DR   N-PSDB; X14598.
XX
PT   New isolated Helicobacter polynucleotides - used to develop products
PT   for the diagnosis, prevention and treatment of Helicobacter
PT   infections and gastrointestinal diseases
XX
PS   Claim 8; Page 2000-2008; 2054pp; English.
XX
CC   This sequence represents a Helicobacter pylori GHPO protein of the
```

```
CC   invention. The polypeptides can be used for preventing or treating
CC   Helicobacter infections, and gastroduodenal diseases associated with
CC   these infections, including acute, chronic, and atrophic gastritis, and
CC   peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC   used for the production of antibodies. The products can also be used for
CC   detection and diagnosis.
XX
SQ   Sequence   1743 AA;


  Query Match             21.8%;  Score 60;  DB 19;  Length 1743;
  Best Local Similarity   34.2%;  Pred. No. 94;
  Matches   13;  Conservative    9;  Mismatches   12;  Indels    4;  Gaps    1;

Qy       7 KRYEDCRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQ 44
           | |:||  :     | :||:|:||:    :  :| :||
Db     621 kaykdcvsk----arnekekqecekllltpearkkleqq 654


RESULT  15
W62835
ID   W62835 standard; Protein; 593 AA.
XX
AC   W62835;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Zea mays antimicrobial protein.
XX
KW   antimicrobial protein; infestation; control.
XX
OS   Zea mays.
XX
PN   WO9827805-A1.
XX
PD   02-JUL-1998.
XX
PF   22-DEC-1997;   97WO-AU00874.
XX
PR   20-DEC-1996;   96AU-0004275.
XX
PA   (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
PI   Bower NI,  Goulter KC,  Green JL,  Manners JM,  Marcus JP;
XX
DR   WPI; 1998-377279/32.
XX
PT   Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT   useful for controlling microbial infestations of plants or mammals
XX
PS   Claim 1; Page 58-60; 96pp; English.
XX
CC   The sequence is that of an antimicrobial protein which can
CC   be used to control microbial infestations in plants and mammalian
CC   animals.
XX
SQ   Sequence   593 AA;


  Query Match             21.6%;  Score 59.5;  DB 19;  Length 593;
  Best Local Similarity   38.9%;  Pred. No. 36;
  Matches   14;  Conservative    7;  Mismatches   12;  Indels    3;  Gaps    2;

Qy      12 CRRRCEWDTRGQKEQQQCEESCKSQYGEKDQQQ-RH 46
           | ||||   |   :: :| | |: :  || |:: ||
Db      40 cvrrce--drpwhqrprcleqcreeerekrqersrh 73


Search completed: March  1, 2001, 15:47:21
Job time: 246 sec
```